APPENDIX

Vol. 33, No. 5 February 1, 1974

A COMPOSITE EFFORT TO BUILD AN ON-LINE NATIONAL SERIALS DATA BASE

by Lawrence G. Livingston

[Editor's Note: The author, Program Officer of the Council on Library Resources, Inc., and Head of the CLR's Systems Group, delivered this paper to the midwinter meeting of the Association of Research Libraries at Chicago on January 19. The paper was distributed to interested American Library Association groups in the following week.]

An urgent requirement exists for a concerted effort to create a comprehensive national serials data base in machine-readable form. Neither the National Serials Data Program nor the MARC Serials Distribution Service, at their current rate of data base building, will solve the problem quickly enough. Because of the absence of a sufficient effort at the national level, several concerted efforts by other groups are under way to construct serials data bases. These institutions have been holding in abeyance the development of their automated serials systems, some for several years, waiting for sufficient development at the national level to provide a base and guidance for the development of their individual and regional systems. This has not been forthcoming, and local pressures from their users, their administrators and their own developing systems are forcing these librarians to act without waiting for the national effort. These efforts are exemplified by the work of one group of librarians, described below.

What has now come to be known as the "Ad Hoc Discussion Group on Serials" had its beginnings in an informal meeting during the American Library Association's Conference in Las Vegas last June. You will also hear this discussion group referred to as the "Toronto Group." This is because its prime mover has been Richard Anable of York University, Toronto, and because the first formal meeting occurred in that city. The expenses of the Toronto and subsequent meetings have been borne by the Council on Library Resources, and Council staff have been involved in each meeting. A fuller exposition of the origins, purposes and plans of the Toronto group has been written by Mr. Anable for the *Journal of Library Automation.* It appeared in the December 1973 issue.

Quoting from Anable: At the meeting (in Las Vegas) there was a great deal of concern expressed about:

1. The lack of communication among the generators of machine-readable serials files,
2. The incompatibility of format and/or bibliographic data among existing files,
3. The apparent confusion about the existing and proposed bibliographic description and format "standards." End of quote.

The Toronto Group agreed that something could and should be done about these problems. If nothing else, better communications among those libraries and systems creating machine-readable files would allow each to enhance its own systems development by taking advantage of what others were doing.

As the discussions progressed, several points of consensus emerged. Among them were:

1. The MARC Serials Distribution Service of the Library of Congress and the National Serials Data Program together were not building a national serials data base in machine-readable form fast enough to satisfy the requirements of developing library systems. This systems development was, in several places, at the point where it could no longer wait on serials data base development at the national level as long as progress remained at the current rate.

2. The MARC serials format developed at LC offered the only hope for machine format capability. Every system represented planned to use it. For the purpose of building a composite data base outside LC, the MARC serials format would probably require minor modification, principally by extension. These extensions could and should be added on so as to do no violence to software already developed to handle MARC serials.

3. There existed some difference between the LC MARC serials format and that used by the National Serials Data Program. These differences arose from several circumstances. For example, the MARC serials format predated the International Serials Data System (ISDS), the National Serials Data Program and the key title concept. When these three came along, the requirement existed that the NSDP abide by the conventions of the ISDS. Since the key title is not yet a cataloging title, but is the title to which the International Standard Serial Number is assigned, it is natural that the approach to serial record creation by NSDP should be different from that of a library cata-

loging serials by conventional methods. A working group under the auspices of the IFLA Cataloguing Secretariat has devised an International Standard Bibliographic Description for Serials. The working group's recommendations are to be distributed for trial, discussion and recommendation for change in February. When the ISBD(S) is accepted into cataloguing practice, some of the differences in MARC usage and NSDP procedure will disappear. Others will still remain and they *must* be reconciled. We cannot continue with two serial records, both of which claim to be national in purpose but which are incompatible with each other. A good exposition of the differences in these serials records from the point of view of the MARC Development Office is in an article by Mrs. Josephine Pulsifer in the December 1973 issue of the *Journal of Library Automation*.

4. Major Canadian libraries are active in cooperative work on serials and these two national efforts should be coordinated.

Several other circumstances bear on the problem. For example, the National Serials Data Program is a national commitment of the three national libraries. In addition to the funding from the three national libraries, there are excellent chances that the NSDP will receive funds from other sources to expedite its activities. The NSDP is responsible for the ISSN and key title and for relationships with the International Serials Data System. Ultimately, the ISSN and key title will be of great importance to serials handling in all libraries. For all of these reasons, it is imperative that the activities of the NSDP be channeled into the comprehensive data base building effort described in this paper.

When it was realized at the Council on Library Resources that the Toronto Group was serious and that a data base building effort would result, it was obvious that this had enormous significance for the Library of Congress and other library systems because the result would be a *de facto* national serials data base. Accordingly, a paper was prepared and sent to LC, urging that an effort be made in Washington to coordinate the efforts of the MARC Serials Distribution Service, the National Serials Data Program and this external effort. In addition, it was felt that LC should take a hard look at its own several serials processing flows and attempt to reconcile them better with each other and with the external effort. To do this, LC was urged to do a brief study of LC serials systems, using LC staff and one person from CLR. LC agreed and the study is now very nearly complete. The written guidance given the study group members was quite specific. They were to study all serials flow at LC and make their recommendations based on what LC should be doing, rather than being constrained by what LC is doing. The overall objectives of the study were to aim for the creation of serials records as near the source as possible and one-time conversion of each record to machine-readable form to serve multiple uses. Specifically to be examined were the serials processing flows of the Copyright Office, the Order Division, the Serial Record Division, New Serials Titles and the National Serials Data Program.

While all of this was going forward, the Toronto Group had some more meetings. OCLC was tentatively selected as the site for the data base building effort. It is understood by everyone that this is a temporary solution; eventually a national-level effort must be mounted which will provide a post-edit capability to bring the composite data base up to nationally acceptable standards. A permanent update capability is also required. This permanent activity, hopefully, will be based at the Library of Congress. OCLC was chosen as the interim site for several reasons, but especially for its proven capability to produce network software and support which will work. Within a very short time OCLC will have on-line serials cataloguing and input capability which will extend to some two hundred libraries. No other system is nearly so far advanced. The Toronto Group has assured itself that the data record OCLC intends to use is adequate and is now working on the conventions required to insure consistency in input and content, to include some recommendations for minor additions to the MARC serials format.

During their deliberations, the Toronto Group realized that, to be effective, their efforts needed formal sponsorship, and discussions to this end were begun. Initially, several agencies were considered to be candidates for this management role. Various considerations quickly narrowed the list down to the Library of Congress, The Association of Research Libraries and the Council on Library Resources, and representatives of these three met to discuss the matter further. During the discussions, CLR was asked to assume the interim management responsibility until a permanent arrangement could be worked out. CLR was selected because, as an operating foundation under the tax laws, it can act expeditiously in matters of this kind. CLR can also deal with all kinds of libraries and has no vested interest in any particular course of action.

Meanwhile, certain institutions in the Toronto

Group had indicated that they were ready to pledge $10,000 among themselves for the specific purpose of hiring Mr. Anable as a consultant to continue his coordinating activities. The group asked CLR to act as agent to collect and disburse these funds.

CLR is ready to assume the initial responsibility for the management of this cooperative data base building effort, if that is the will of the leadership in the library community. CLR is prepared to commit one staff member full time to the project who is well versed in the machine handling of MARC serials records. This is Mr. George Parsons, and other staff members will assist as appropriate. Mr. Anable has agreed to act as a consultant to help coordinate these activities. CLR would aim for the most complete, accurate and consistent serial record in the LC MARC serials format which can be had under the circumstances. During the effort, CLR will act as the point of contact between OCLC and the participating libraries, assisting in negotiating contracts and other agreements as required. The composite data base will be made available to all other libraries as the least possible cost for copying. Initially at least, the costs of this effort will have to be shared by the participating libraries, since no additional funds are presently available. The goal is to build 100,000 serial records the first year, another 100,000 the second year, and design and implement the permanent mechanism the third year, while file-building continues.

As the project gets under way, it will work like this: a set of detailed written guidelines for establishing the record and creating the input will be promulgated, and agreement to abide by them will be a prerequisite to participation. Selected libraries with known excellence in serial records will be asked to participate; others may request participation. Those selected who already have or can arrange for terminals on the OCLC system will participate on line. This is the preferred method, but it may be possible to permit record creation off line, such records to be added to the data base in a batch mode. It is very difficult to merge serial files from different sources in this way, so an attempt will be made to find a large serials data base in machine-readable form for use as a starting point. This file would be read into the OCLC system. A participating library wishing to enter a record would first search to see whether it existed in the initial data base. If a record is found, it would be updated insofar as this is possible, within the standards chosen for the system. It may be further updated by other participants, still within the system standards, but at some point update on a record in the system will reach a point of diminishing returns and the record will remain static until a post-edit at the national level can be performed. These records will be for use as their recipients see fit, but their prime purpose is to support the development of automated serials systems while eliminating duplication of effort.

Details of how to flag these records in the OCLC data base as they are being created by this effort will be worked out, as will be the relationship between this effort and the rest of OCLC activities. CLR will, from time to time, report progress to the community.

It would be the hope of CLR that the Toronto Group will continue to assist in the technical and detailed aspects of the project. In addition, and after consultation with the appropriate people, an advisory group will be appointed to advise CLR in this effort.

Ad Hoc Discussion Group on Serials Data Bases
(The "Toronto Group")*

Richard Anable
York University
Downsview, Ontario, Canada

Henriette Avram
Chief, MARC Development Office
Library of Congress
Washington, D.C. 20540

Isabel Bradley
International Serials Data System
National Library of Canada
Ottawa, Ontario, Canada

Jay L. Cunningham, Director
University-wide Library Automation Program
Institute of Library Research
University of California
South Hall Annex
Berkeley, California 94720

Rosario de Varennes
Bibliothèque Generale
Université Laval
Quebec, Quebec GIK 7P4
Canada

*Some did not attend all meetings.

R. M. Duchesne (observer)
Systems Development Branch/
British Library
London, England

Glyn T. Evans
State University of New York
99 Washington Ave.
Albany, New York 12210

Audrey N. Grosch
University of Minnesota
Minneapolis, Minnesota 55455

Alma Harvey
Head, Cataloguing and Classification
National Science Library
National Research Council of Canada
100 Sussex Drive, Room 2008
Ottawa, Ontario, Canada

Karl F. Heumann
Chairman, Joint Committee on Union List of Serials
Federation of American Societies for Experimental Biology
9650 Rockville Pike
Bethesda, Maryland 20014

Joseph H. Howard
Chief, Serial Record Division
Library of Congress
Washington, D.C. 20540

Joseph Judy
Room 002
National Agricultural Library
Beltsville, Maryland 20705

Lawrence G. Livingston
Council on Library Resources
Suite 620, One Dupont Circle, N.W.
Washington, D.C. 20036

Robert Lynch
University of Massachusetts
Amherst, Massachusetts 01002

John McGowan
Association of Research Libraries/
Ad Hoc Committee on Machine-Based Serial Records
Northwestern University
Evanston, Illinois 60210

Ron Miller
NELINET/New England Board of Higher Education
40 Grove Street
Wellesley, Massachusetts 02181

George Parsons
Council on Library Resources
Suite 620, One Dupont Circle, N.W.
Washington, D.C. 20036

Flora Patterson
National Library of Canada
Ottawa, Ontario, Canada K1A ON4

Joseph W. Price
National Serials Data Program
Room A5121
Library of Congress
Washington, D.C. 20540

Cynthia Pugsley
Assistant Librarian, Planning
University of Waterloo Library
Waterloo, Ontario, Canada N2L 3G1

Basil Stuart-Stubbs
Canadian Union Catalog Task Group
University of British Columbia
British Columbia, Canada

James Rush
Ohio College Library Center
1550 W. Henderson
Columbus, Ohio 43220

Paul Vassallo
National Serials Data Program
Library of Congress
Washington, D.C. 20540

Velma Veneziano
MARBI Committee
Northwestern University
Evanston, Illinois 60201